

# SEQUENCE LISTING

<110> Genencor International, Inc.

Jones, Brian E.

Grant, William D.

Heaphy, Shaun

Grant, Susan

<120> Novel Bacillus BagCel Cellulase

<130> GC798-2-PCT

<140> PCT/US2004/013175

<141> 2004-04-28

<150> US 60/467,255

<151> 2003-04-30

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4205

<212> DNA

<213> Bacillus sp.

<220>

<221> misc\_feature

<222> (1)...(4205)

<223> isolated from environmental water sample from  
Sonachi Lake, Kenya

<400> 1

tctcatttag	tttgactgaa	atacctactg	tgccacaaag	taaagttaaa	ctgacgaatg	60
tggaacgaat	cacttaatgg	ttctagcata	gataacgaaa	gatgaacacg	ttcaaagtgc	120
gccactcttt	tgaagagagg	tgaacttttt	ttgtgacaag	aaaggggtgt	aaatgaagat	180
caaagctaaa	caagatgagt	aacgtttctt	ttctttttta	tagagtgagt	tagtatatga	240
gtcccttata	aattttctaga	ctgttatatt	aaataattga	atgactcagt	caccattaag	300
ttttcaacac	ccataagcga	cgtttgaaga	tctaatagat	cgagagggtt	tatcactttg	360
gagcggaaga	tactgttagg	actcgtttta	tatggtgaac	ttggtgttat	tgtgtattta	420
aaagaaagg	aaacgaaaaa	aatgggttaa	ttagaaagag	gctattacag	agaggagAAC	480
aaacaatgaa	cgtaacactt	gaagtgcac	actgcacgac	taaagggtat	cgaacaacct	540
ttcattcaga	aggatatggg	gccgaaaaag	caattaccat	cgcagaagat	tttcagcggg	600
caggacggat	aaaacagatc	gtcttttagg	atgagcgtga	tagtccgtgg	acgttaaaag	660
aaacttaaaag	attttttagaa	gagattaaaa	cggagccgca	tcatctctct	gtgtattttg	720
atgggggatt	tgattttggg	acacaacgat	ctgggtcttg	gtgtgattta	ttatgaacaa	780
aatgacacgt	cttatcgggt	gagaagaaac	gctaccgtgg	cgtcattgac	atcgaataac	840
gaagcagaat	atgccgcttt	acatttagga	cttaaaagAAC	ttgaagggat	cgggtgcgcat	900
catctaccta	tactatttta	cgggtgattct	caagttgtga	tcaatcagtt	aaaaggagaa	960
tgggcgtgta	tggaggagg	gttaaataaa	tgggctgacc	gtattgatca	gcatttagct	1020
aaattaggca	tgaccgctac	ttataagtta	atcccccgta	aagaaaaccg	tgaagcagat	1080
caactggcta	cacaagcgtt	aaacgggcaa	gaaattataa	gtcaacgtga	tgtcagtgag	1140
cgtggtgcag	attagtctgc	acccgcataa	aagttaacgt	atatagaagt	ggatggggat	1200

taaaggaacg	tcattcactc	taagcaagcg	ttgcgacagc	aaaaaagaaa	catataaggt	1260
ttttctgagc	tactatctat	acaaatagcc	aagtggcagt	taagctctta	cctcatcaag	1320
tttttgacta	ccagtcttcc	actcctactt	tcacctatat	aaattgggtc	cttttttgtt	1380
aataatcact	aattttgacg	gtatttttta	atagaaatat	atgctagatt	ataaactagt	1440
aacgatgtag	aaggtgggtg	ttgaccatat	aagaagactc	tttcaaacct	ggtagtatcg	1500
cattaaaaaa	tttgaaagggt	ggagaggaca	catgggttat	accaaagcga	agtgtacgtt	1560
gaaaaaaact	gtcttgtttg	gtttaattct	ctgtttaagt	gtgtcaatgt	ttgttccaat	1620
gacatcagct	gaagatgtca	cttcgtcaca	gttgatatt	cactcctatg	tagctgacat	1680
gcagcctggc	tggaaatttag	gaaatacggt	tgacgctgtt	ggagatgatg	aaacagcgtg	1740
ggggaatcct	cgtgtaacaa	gagagttaat	aaaaacgatt	gctgatgaag	ggataaaaag	1800
cattcgtatc	ccagtgcacat	ggcaaaatca	aatgggtggg	tctccagatt	atacgataaa	1860
tgaagattat	atcaatcggg	tgagcaagc	gatagattgg	gcgttggagg	aagacttata	1920
tgtgatgtta	aatgtgcatc	atgactcatg	gctgtggatg	tatgatatgg	aacataacta	1980
tgatgaggtc	atggcaagat	atacagctat	ttgggaacaa	ttgtcggaaa	aattcaaaaag	2040
ccactcccat	aagttgatgt	ttgagagtgt	caatgagcct	aggtttacgc	aggagtgggg	2100
agagattcaa	gaaaatcatc	atgcttactt	agaagattta	aataagacgt	tctattatat	2160
tgtcagagag	tcaggaggca	ataatgtgga	gcgcccttta	gtattgccta	cgatagaaac	2220
agccacgtct	caggatttac	tagatcgctt	gtatcaaaca	atggaagact	tggatgatcc	2280
ttatttaatt	gccacggtgc	attattatgg	cttctggcca	tttagtgtca	atatagcagg	2340
gtacactcat	tttgaacagg	aaacacaaca	agatattata	gacacctttg	accgtgttca	2400
taacacattt	acagcgcgtg	gtgtcccagt	tgtattaggc	gaattcgggt	tgttaggctt	2460
tgacaaaagt	acggatgtga	ttcagcaagg	ggagaaatta	aagttttttg	agtttctcat	2520
ccatcatctc	aatgaacgtg	atataaccca	tatgttatgg	gataacggcc	agcattttaa	2580
tcgagaaaact	tatgcatggt	atgatcaaga	atttcatgac	atattaaaag	cgagttggga	2640
ggggcgttct	gctacagcag	agtctaattt	gattcatgtg	aaggacggaa	agccaattag	2700
agatcaagat	atacagcttt	acttaaacgg	aaatgagcta	acagccttac	aggcagggga	2760
ggaatcgctt	gttctaggag	aggattatga	actagcagga	ggcgtattaa	cgctaaaagc	2820
ggacaccctc	acaagactaa	ttacccttgg	tcaattagga	accaatgcag	tcacacagc	2880
acaatttaat	tctggagcag	actggcggtt	tcaattacag	aatgtggacg	tgccaacggt	2940
cgaaaaataca	gatggctcaa	catggcattt	tgcgatccct	acccatttta	atgggtgatag	3000
tcttgcgacg	atggaagctg	tttatgcaaa	cggagaatat	gctggggccgc	aagattggac	3060
gtcatttaaa	gaatttggcg	aggcggtttc	tcctaattac	gccacagggg	aaattattat	3120
atcagaagcc	ttctttaacg	cggtagcgga	tgatgatatc	catttaacat	ttcatttttg	3180
gagcggagag	acggtggaat	ataccttacg	taaaaatggc	aattatgttc	aaggtagacg	3240
gtaacatgat	tttaattaat	agataaacca	gcctacctat	cgtttttgga	agaaggcaaa	3300
cgaatctcat	cttaccaaca	ccgtgcttta	gaactttaga	agtgcggtg	ttttttaaga	3360
catgaggaga	agcaatcctc	tatcaacagt	caccaatttt	tattcaggag	gtgtcaagtt	3420
atctaacgtt	ctatgaatgc	atatagtctc	tgacgaataa	acatagttaa	aaagaagtga	3480
gcctagtccc	cgagggggaag	gggataatgc	caacgtattg	gattaaagta	ccttcttgat	3540
aaaaagaaaag	ggttttcaag	agatggaaat	gggctcggtt	gttatacttt	aattacgcct	3600
tggaacgtca	ttttggcggt	gctatttagt	aagatggctg	acatcataaa	agaggagtgg	3660
gttcgatggc	tttaattcaa	ttaagcttta	aatcacgagc	attaatgttg	caaacctctg	3720
tcaatgtttt	attaccgggtg	ggaatgaatg	ctgtagattt	tacaccaagt	gatgattttt	3780
cttatgttac	tgaccctttt	cctgtcctat	atcttttgca	tggtgcaact	gatgattatt	3840
cagcatggct	acgtctgtcc	tctattgaac	gatatgctga	agaaaaaaaa	ttggcggtcg	3900
tcatgccaaa	tgctgatatg	agtgcgtata	cggatatggt	acatggccat	cgttactgga	3960
cgtatattag	taaggtgctg	cctgagttta	tgagagcaac	ttttcctatt	tctcagcacc	4020
gtgaagacac	ctttgcagct	ggtctgtcta	tgggaggata	cggggctttt	aaattggcgc	4080
tgcggaacc	ggaacgcttc	gctgcagctg	tgtcattatc	tggtgcagtt	gatatgagag	4140
aagcaagtca	accagactcc	ctatttghta	atgcctttgg	tgaagggacg	aaaatcgcag	4200
ggaca						4205

<210> 2

<211> 1713

<212> DNA

<213> Bacillus sp.

<220>  
 <221> misc\_feature  
 <222> (1)...(1713)  
 <223> isolated from environmental water sample from  
 Sonachi Lake, Kenya

<400> 2

atgggttata	ccaaagcgaa	gtgtacgttg	aaaaaaactg	tcttgtttgg	tttaattctc	60
tgtttaagt	tgtcaatggt	tggtccaatg	acatcagctg	aagatgtcac	ttcgtcacag	120
ttggatattc	actcctatgt	agctgacatg	cagcctggct	ggaatttagg	aaatacgttt	180
gacgctgttg	gagatgatga	aacagcgtgg	gggaatcctc	gtgtaacaag	agagttaata	240
aaaacgattg	ctgatgaagg	gtataaaagc	attcgtatcc	cagtgacatg	gcaaaatcaa	300
atgggtgggt	ctccagatta	tacgataaat	gaagattata	tcaatcgggt	ggagcaagcg	360
atagattggg	cggtggagga	agacttatat	gtgatgttaa	atgtgcatca	tgactcatgg	420
ctgtggatgt	atgatatgga	acataactat	gatgagggtca	tggcaagata	tacagctatt	480
tgggaacaat	tgtcggaaaa	attcaaaagc	cactcccata	agttgatgtt	tgagagtgtc	540
aatgagccta	ggtttacgca	ggagtgggga	gagattcaag	aaaatcatca	tgcttactta	600
gaagatttaa	ataagacgtt	ctattatatt	gtcagagagt	caggaggcaa	taatgtggag	660
cgccctttag	tattgcctac	gatagaaaca	gccacgtctc	aggatttact	agatcgcttg	720
tatcaaacaa	tggaagactt	ggatgatcct	tatttaattg	ccacggtgca	ttattatggc	780
ttctggccat	ttagtgtcaa	tatagcaggg	tacactcatt	ttgaacagga	aacacaacaa	840
gatattatag	acacctttga	ccgtgttcat	aacacattta	cagcgcgtgg	tgtcccagtt	900
gtattaggcg	aattcggttt	gttaggcttt	gacaaaaagta	cggatgtgat	tcagcaaggg	960
gagaaattaa	agttttttga	gtttctcatc	catcatctca	atgaacgtga	tataacccat	1020
atgttatggg	ataacggcca	gcatttaa	cgagaaactt	atgcatggta	tgatcaagaa	1080
tttcatgaca	tattaaaagc	gagttgggag	gggcgttctg	ctacagcaga	gtctaatttg	1140
attcatgtga	aggacggaaa	gccaattaga	gatcaagata	tacagcttta	cttaaacgga	1200
aatgagctaa	cagccttaca	ggcaggggag	gaatcgcttg	ttctaggaga	ggattatgaa	1260
ctagcaggag	gcgtattaac	gctaaaaagc	gacaccctca	caagactaat	tacccttggt	1320
caattagcaa	ccaatgcagt	catcacagca	caatttaatt	ctggagcaga	ctggcgtttt	1380
caattacaga	atgtggacgt	gccaacggtc	gaaaatacag	atggctcaac	atggcatttt	1440
gcgatcccta	cccattttta	tggtgatagt	cttgcgacga	tggaaagctgt	ttatgcaaac	1500
ggagaatatg	ctgggccgca	agattggacg	tcatttaaag	aatttggcga	ggcgttttct	1560
cctaattacg	ccacagggga	aattattata	tcagaagcct	tctttaacgc	ggtacgggat	1620
gatgatatcc	atttaacatt	tcatttttgg	agcggagaga	cgggtggaata	taccttacgt	1680
aaaaatggca	attatgttca	aggtagacgg	taa			1713

<210> 3  
 <211> 570  
 <212> PRT  
 <213> Bacillus sp.

<220>  
 <221> VARIANT  
 <222> (1)...(570)  
 <223> isolated from environmental water sample from  
 Sonachi Lake, Kenya

<400> 3

Met	Gly	Tyr	Thr	Lys	Ala	Lys	Cys	Thr	Leu	Lys	Lys	Thr	Val	Leu	Phe
1				5					10					15	
Gly	Leu	Ile	Leu	Cys	Leu	Ser	Val	Ser	Met	Phe	Val	Pro	Met	Thr	Ser
			20					25					30		
Ala	Glu	Asp	Val	Thr	Ser	Ser	Gln	Leu	Asp	Ile	His	Ser	Tyr	Val	Ala
			35					40						45	

Asp	Met	Gln	Pro	Gly	Trp	Asn	Leu	Gly	Asn	Thr	Phe	Asp	Ala	Val	Gly
50						55					60				
Asp	Asp	Glu	Thr	Ala	Trp	Gly	Asn	Pro	Arg	Val	Thr	Arg	Glu	Leu	Ile
65					70					75					80
Lys	Thr	Ile	Ala	Asp	Glu	Gly	Tyr	Lys	Ser	Ile	Arg	Ile	Pro	Val	Thr
			85						90					95	
Trp	Gln	Asn	Gln	Met	Gly	Gly	Ser	Pro	Asp	Tyr	Thr	Ile	Asn	Glu	Asp
			100					105					110		
Tyr	Ile	Asn	Arg	Val	Glu	Gln	Ala	Ile	Asp	Trp	Ala	Leu	Glu	Glu	Asp
		115						120					125		
Leu	Tyr	Val	Met	Leu	Asn	Val	His	His	Asp	Ser	Trp	Leu	Trp	Met	Tyr
130						135					140				
Asp	Met	Glu	His	Asn	Tyr	Asp	Glu	Val	Met	Ala	Arg	Tyr	Thr	Ala	Ile
145					150					155					160
Trp	Glu	Gln	Leu	Ser	Glu	Lys	Phe	Lys	Ser	His	Ser	His	Lys	Leu	Met
				165					170					175	
Phe	Glu	Ser	Val	Asn	Glu	Pro	Arg	Phe	Thr	Gln	Glu	Trp	Gly	Glu	Ile
			180					185					190		
Gln	Glu	Asn	His	His	Ala	Tyr	Leu	Glu	Asp	Leu	Asn	Lys	Thr	Phe	Tyr
		195					200					205			
Tyr	Ile	Val	Arg	Glu	Ser	Gly	Gly	Asn	Asn	Val	Glu	Arg	Pro	Leu	Val
210						215					220				
Leu	Pro	Thr	Ile	Glu	Thr	Ala	Thr	Ser	Gln	Asp	Leu	Leu	Asp	Arg	Leu
225					230					235					240
Tyr	Gln	Thr	Met	Glu	Asp	Leu	Asp	Asp	Pro	Tyr	Leu	Ile	Ala	Thr	Val
			245						250					255	
His	Tyr	Tyr	Gly	Phe	Trp	Pro	Phe	Ser	Val	Asn	Ile	Ala	Gly	Tyr	Thr
			260					265					270		
His	Phe	Glu	Gln	Glu	Thr	Gln	Gln	Asp	Ile	Ile	Asp	Thr	Phe	Asp	Arg
		275					280					285			
Val	His	Asn	Thr	Phe	Thr	Ala	Arg	Gly	Val	Pro	Val	Val	Leu	Gly	Glu
290						295					300				
Phe	Gly	Leu	Leu	Gly	Phe	Asp	Lys	Ser	Thr	Asp	Val	Ile	Gln	Gln	Gly
305					310					315					320
Glu	Lys	Leu	Lys	Phe	Phe	Glu	Phe	Leu	Ile	His	His	Leu	Asn	Glu	Arg
				325					330					335	
Asp	Ile	Thr	His	Met	Leu	Trp	Asp	Asn	Gly	Gln	His	Leu	Asn	Arg	Glu
			340					345					350		
Thr	Tyr	Ala	Trp	Tyr	Asp	Gln	Glu	Phe	His	Asp	Ile	Leu	Lys	Ala	Ser
		355				360						365			
Trp	Glu	Gly	Arg	Ser	Ala	Thr	Ala	Glu	Ser	Asn	Leu	Ile	His	Val	Lys
370						375					380				
Asp	Gly	Lys	Pro	Ile	Arg	Asp	Gln	Asp	Ile	Gln	Leu	Tyr	Leu	Asn	Gly
385					390					395					400
Asn	Glu	Leu	Thr	Ala	Leu	Gln	Ala	Gly	Glu	Glu	Ser	Leu	Val	Leu	Gly
			405						410					415	
Glu	Asp	Tyr	Glu	Leu	Ala	Gly	Gly	Val	Leu	Thr	Leu	Lys	Ala	Asp	Thr
			420					425					430		
Leu	Thr	Arg	Leu	Ile	Thr	Pro	Gly	Gln	Leu	Gly	Thr	Asn	Ala	Val	Ile
		435					440					445			
Thr	Ala	Gln	Phe	Asn	Ser	Gly	Ala	Asp	Trp	Arg	Phe	Gln	Leu	Gln	Asn
450						455					460				
Val	Asp	Val	Pro	Thr	Val	Glu	Asn	Thr	Asp	Gly	Ser	Thr	Trp	His	Phe
465					470					475					480
Ala	Ile	Pro	Thr	His	Phe	Asn	Gly	Asp	Ser	Leu	Ala	Thr	Met	Glu	Ala
			485					490					495		
Val	Tyr	Ala	Asn	Gly	Glu	Tyr	Ala	Gly	Pro	Gln	Asp	Trp	Thr	Ser	Phe

			500						505				510				
Lys	Glu	Phe	Gly	Glu	Ala	Phe	Ser	Pro	Asn	Tyr	Ala	Thr	Gly	Glu	Ile		
		515					520					525					
Ile	Ile	Ser	Glu	Ala	Phe	Phe	Asn	Ala	Val	Arg	Asp	Asp	Asp	Ile	His		
	530					535					540						
Leu	Thr	Phe	His	Phe	Trp	Ser	Gly	Glu	Thr	Val	Glu	Tyr	Thr	Leu	Arg		
545					550					555					560		
Lys	Asn	Gly	Asn	Tyr	Val	Gln	Gly	Arg	Arg								
			565					570									